

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-228473

(43)Date of publication of application : 25.08.1998

(51)Int.Cl. G06F 17/24
G06F 17/27
G06F 17/21
G06T 11/60

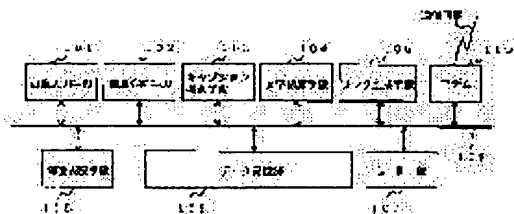
(21)Application number : 09-028955 (71)Applicant : RICOH CO LTD
(22)Date of filing : 13.02.1997 (72)Inventor : SAITO TAKASHI

(54) DOCUMENT PICTURE PROCESSING METHOD, DOCUMENT PICTURE PROCESSOR AND STORAGE MEDIUM

(57)Abstract:

PROBLEM TO BE SOLVED: To automatically generate a link among a drawing, a list and a text related to them although a caption exists or not and to make it into a hyper text.

SOLUTION: An area division means 102 divides a document picture into the elements of a character area, the drawing and the list. A caption discrimination means 103 discriminates the character area equivalent to the caption of the drawing and the list. A character recognition means 104 recognizes the character area and a link generation means 105 generates the link among the list, the caption and the pertinent place of the text related to them. A structure expression means 106 expresses link information and the like by HTML(hyper text mark up language).



特開平 1 0 - 2 2 8 4 7 3

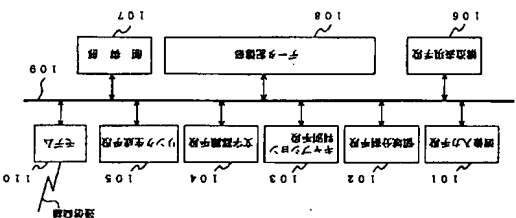
(43) 公開日 平成10年(1998)8月25日

(51) Int. Cl. ⁶	機別記号	F I
G 0 6 F	17/24	G 0 6 F
	17/27	15/20
	17/21	5 4 6 A
		5 5 0 E
G 0 6 T	11/60	5 7 0 D
		15/62
		3 2 5 K
審査請求	未請求	請求項の数 7
		O L
		(全 6 頁)

(21) 出願番号	特願平9-28955	(71) 出願人	000006747
			株式会社リコー
(22) 出願日	平成9年(1997)2月13日		東京都大田区中馬込丁目3番6号
		(72) 発明者	齋藤 清志
			東京都大田区中馬込丁目3番6号
			株式会社リコー内
		(74) 代理人	弁理士 結木 誠 (外1名)

(54) 【発明の名称】 文書画像処理方法、文書画像処理装置および記憶媒体

(57) 【要約】
【課題】 キヤブジョンがあるなしに係らず、図、表とそれに関連する本文との間にリンクを自動的に生成し、ハイパーテキスト化する。
【解決手段】 領域分割手段102は文書画像を文字領域、図、表などの要素に分割し、キヤブジョン判別手段103は図、表などのキヤブジョンに相当する文字領域を判別する。文字認識手段104は文字領域を認識し、リンク生成手段105は図、表やキヤブジョンと、それに関連する本文の該当箇所の間にリンクを生成し、構造表現手段106によってリンク情報などをHTMLで表現する。



【特許請求の範囲】

【請求項1】 文書画像を、文字領域、図、表（以下、図表）領域を含む要素に分割し、該図表領域とその近傍の文字領域との位置関係を基に該図表に関連するキヤブジョンがあるか否かを検出し、該キヤブジョンが検出されるとき、該キヤブジョン内から前記図表に関連する所定の文字列を検出し、該検出された文字列と同一の文字列を前記文字領域から検出し、前記キヤブジョン内の文字列と前記文字領域内の文字列との間にリンクを生成することを特徴とする文書画像処理方法。

【請求項2】 前記図表に関連する所定の文字列は、図表の番号であることを特徴とする請求項1記載の文書画像処理方法。

【請求項3】 文書画像を、文字領域、図表領域を含む要素に分割し、該図表領域とその近傍の文字領域との位置関係を基に該図表に関連するキヤブジョンがあるか否かを検出し、該キヤブジョンが検出されるとき、前記図表に関連する所定の文字列を、前記文字領域から検出し、該文字領域内の文字列と前記図表との間にリンクを生成することを特徴とする文書画像処理方法。

【請求項4】 前記図表に関連する所定の文字列は、レイアウトに関する表現であることを特徴とする請求項2記載の文書画像処理方法。

【請求項5】 前記リンク生成された文書画像をハイパーテキスト化することを特徴とする請求項1または2記載の文書画像処理方法。

【請求項6】 文書画像を入力する手段と、該入力された文書画像を、文字領域、図表領域を含む所定の領域に分割する手段と、該文字領域内の文字を認識する手段と、前記図表と文字との間に、または文字間にリンクを生成する手段と、該リンク生成された文書画像をハイパーテキスト化する手段と、該生成したリンクを用いてハイパーテキストを表示する手段とを備えたことを特徴とする文書画像処理装置。

【請求項7】 文書画像を、文字領域、図表領域を含む要素に分割する機能と、該図表領域とその近傍の文字領域との位置関係を基に該図表に関連するキヤブジョンがあるか否かを検出する機能と、該キヤブジョンが検出されるとき、該キヤブジョン内から前記図表に関連する所定の文字列を検出する機能と、該検出された文字列と同一の文字列を前記文字領域から検出する機能と、前記キヤブジョン内の文字列と前記文字領域内の文字列との間にリンクを生成する機能と、前記キヤブジョンが検出されるとき、前記図表に関連する所定の文字列を、前記文字領域から検出する機能と、該文字領域内の文字列と前記図表との間にリンクを生成する機能と、該リンク生成された文書画像をハイパーテキスト化する機能と、該生成したリンクを用いてハイパーテキストを表示する機能をコンピュータに実現させるためのプログラムを記憶した記憶媒体。

(2) 特開平 1 0 - 2 2 8 4 7 3

【発明の詳細な説明】

【0001】
【発明の属する技術分野】 本発明は、文書画像からハイパーテキスト（HTML文書）を自動的に生成する文書画像処理方法、文書画像処理装置および記憶媒体に関する。

【0002】
従来の技術1入力した文書画像を有効に利用するためには、文書画像のままでは扱いにくいので、レイアウト解析技術やOCR等を利用して構造化文書にするのが望ましい。HTML（Hyper Text Markup Language）文書はその一例であり、文書内または文書間にリンク関係を表現することができる。

【0003】 文書をハイパーテキスト化する従来技術として、例えば、第1の公知例（特開平7-98708号公報）に記載された「文書処理システムおよびその方法」では、読み手の意図に従って作成済みの文書を読み手に適合するように再構成して出力し、またハイパー文書を構成することにより文書の参照方法を任意に決定できるシステムがある。

【0004】 また、第2の公知例として、藤井他「イメージを対象としたHTML自動生成方式の検討」電子通信学会 信学技報 OSE 95-23、1E95-55（1995-09）では、文書画像（複色ページ）からなるイメージとして入力し、それをHTML文書として扱えるように各ページイメージ毎に与えた属性からHTML文書を自動生成する方法がある。

【0005】 さらに、第3の公知例として、大関他「電子図書館構築のための印刷文書のハイパーテキスト化」画像の認識・理解シンポジウム（NIRU'96）1996年7月、および山田他「適用性を高めた印刷文書の構造化処理の手法」電子情報通信学会1995年総大会 D-581がある。これら2つの論文に記載されている手法では、まず抽出した文字領域と図表領域をコラム位置の推測によって文字ブロックと図表ブロックに識別し、図表ブロックに含まれているキヤブジョンの文字認識結果により図表番号をラベルとして与える。そして本文中の参照箇所（「図11」などの出典）を検出し、HTML形式でリンク表現を生成する。

【0006】
【発明が解決しようとする課題】 しかし、上記した第1の例では、読み手の意図、レベル、状況に合わせて文書を再構成するために、文書を部品化、構造化、ハイパーテキスト化しているが、特に、図と本文との間にリンクが生成されていない。また、入力文書画像からの当該リンクの自動生成についても記載されておらず、当該リンクを生成することはできない。
【0007】 また、第2の例では、生成したHTML文書はページイメージを単位としており、ページ内の図と本文の間のリンクなどが生成されない。

3

【0008】さらに、第3の例では、キーワードを検出するためにはコラムの存在が必要であり、またキーワードが存在しない場合には図表などと関連する本文の間にリンクを生成することができない。

【0009】本発明の目的は、キーワードがあるなしに係らず、図、表とそれに関連する本文との間にリンクを自動的に生成し、ハイパーテキスト化する文書画像処理方法、文書画像処理装置および記憶媒体を提供することにある。

10

【0010】問題を解決するための手段】本発明を達成するために、請求項1記載の発明では、文書画像を、文字領域、図、表（以下、図表）領域を含む要素に分割し、図表領域とその他の近傍の文字領域との位置関係を基に図表に関連するキーワードがあるか否かを検出し、該キーワードが検出されたとき、該キーワード内から前記図表に関連する所定の文字列を検出し、該検出された文字列と同一の文字列を前記文字領域から検出し、前記キーワード内の文字列と前記文字領域内の文字列との間にリンクを生成することを特徴としている。

【0011】請求項2記載の発明では、前記図表に関連する所定の文字列は、図表の番号であることを特徴としている。

【0012】請求項3記載の発明では、文書画像を、文字領域、図表領域を含む要素に分割し、図表領域とその他の近傍の文字領域との位置関係を基に図表に関連するキーワードがあるか否かを検出し、該キーワードが検出されないとき、前記図表に関連する所定の文字列を検出し、前記文字領域から検出し、該文字領域内の文字列と前記図表との間にリンクを生成することを特徴としている。

【0013】請求項4記載の発明では、前記図表に関連する所定の文字列は、レイアウトに関する表現であることを特徴としている。

【0014】請求項5記載の発明では、前記リンク生成された文書画像をハイパーテキスト化することを特徴としている。

【0015】請求項6記載の発明では、文書画像を入力する手段と、該入力された文書画像を、文字領域、図表領域を含む所定の領域に分割する手段と、該文字領域内の文字を認識する手段と、前記図表と文字との間に、または文字間にリンクを生成する手段と、該リンク生成された文書画像をハイパーテキスト化する手段と、該生成したリンクを用いてハイパーテキストを表示する手段とを備えたことを特徴としている。

【0016】請求項7記載の発明では、文書画像を、文字領域、図表領域を含む要素に分割する機能と、該図表領域とその他の近傍の文字領域との位置関係を基に図表に関連するキーワードがあるか否かを検出する機能と、該キーワードが検出されたとき、該キーワード内か

4

ら前記図表に関連する所定の文字列を検出する機能と、該検出された文字列と同一の文字列を前記文字領域から検出する機能と、前記キーワード内の文字列と前記文字領域内の文字列との間にリンクを生成する機能と、前記キーワードが検出されないとき、前記図表に関連する所定の文字列を、前記文字領域から検出する機能と、該文字領域内の文字列と前記図表との間にリンクを生成する機能と、該リンク生成された文書画像をハイパーテキスト化する機能と、該生成したリンクを用いてハイパーテキストを表示する機能をコンピュータに実現させるためのプログラムを記憶したことを特徴としている。

20

【0017】【発明の実施の形態】以下、本発明の実施例を図面を用いて具体的に説明する。図1は、本発明の実施例の構成を示す。図において、101は画像の入力手段、102は文書画像を文字領域などの要素に分割する領域分割手段、103は図表などのキーワードに相当する文字領域を判別するキーワード判別手段、104は文字行の抽出処理、文字の切り出し処理を含む文字領域の文字を認識する文字認識手段、105は図表やキーワードと、それに関連する本文の該当箇所との間にリンクを生成するリンク生成手段、106はこれらリンク情報やキーワード情報など抽出した情報をHTMLなどで表現する構造表現手段、107は全体の制御部、108は入力された画像データ、領域データ、文字データなどをデータベースを記憶するデータベース部、109はデータ通路、110は適宜回帰に接続されたモデムである。

【0018】図2は、本発明の処理フローチャートを示す。以下、図2を参照しながら本発明の処理動作を説明する。まず、画像入力手段101によって文書画像を得る（ステップ201）。この画像入力手段はスキャナ、フロッパスなどであり、またネットワーク経由で別の装置から画像を得るようにしてもよい。

【0019】次に、入力された文書画像を、領域分割手段102によって文字領域と図、表、図表領域に分割する（ステップ202）。この領域分割方法としては、例えば特開平6-20092号公報に記載された公知技術を用い、例えば、抽出した領域は、属性として図表の種類（文字領域、図、表など）、その位置（領域の外接矩形など）、領域番号などを持つ。

【0020】領域情報が出されると、キーワード判別手段103はキーワードの判別を行う（ステップ203）。図3は、ステップ203のキーワード判別処理の詳細なフローチャートを示す。また、図4は、文書画像が領域分割された例を示す。図において、各領域401、402、404、405は文字領域であり、403は図または表である。以下、図4を例に、キーワード判別処理を説明する。図、表、または図、表以外でキーワードを持つものを処理対象とする（ステップ302）。図4の例では、図または表の領域403が処理対

5

象となる。

【0021】各処理対象毎に以下の処理を行う。まず、キーワード候補となる文字領域を検出する（ステップ303）。これは、文字領域の長さ、処理対象となる図または表との位置関係などの特徴量を基に検出する。図5は、ステップ303の詳細なフローチャートを示す。

例えば、文字領域の長さの絶対値が所定の閾値以内にあること、および/または、文字領域の高さとページ全体の横断文字サイズとの比が所定の閾値以内にあること（ステップ501）、また、位置関係としては図、表領域403と文字領域404との距離406が所定の閾値以内にあること（ステップ502）、図、表領域403と文字領域404との左右（または上）のずれ407、408などが所定の閾値以内であること（ステップ503）が、キーワード候補となる条件であるとする。なお、図4の例では、図表の上にあるキーワードを例にしているが、左右にある場合も同様である。

【0022】図3のフローチャートに戻り、次に、近傍文字領域との位置関係を調べる（ステップ304）。図6は、ステップ304の詳細なフローチャートを示す。図4では、キーワード候補404の近傍に文字領域405が存在する（ステップ601参照）。この2つの文字領域が一つの文章を構成するものである場合は、キーワード候補404が図、表領域403に近い位置にある場合、キーワードとならない。従って、以下のステップではその関係を調べる。

【0023】すなわち、キーワード候補404と近傍文字領域405との距離409、すなわち10、411を調べ、いづれか1つないし複数の値が所定の閾値より大きい場合（ステップ602参照）、また距離409が距離406に比べて十分に大きい場合（ステップ603参照）には、キーワードと判定し、この近傍文字領域405は無視してよい。そうでない場合は（ステップ603参照）、キーワード候補404をキーワードと判定しない。

【0024】文字領域402がキーワード候補の位置にあるが、近傍文字領域401と文字領域402の位置関係が、ステップ602、603の条件を満たすので、文字領域402を非キーワードと判定する。

【0025】以上の処理を繰り返し、対象領域全てについて処理を終えたらこの処理が終了する（ステップ301）。上記したように判定された各キーワード領域は、どの対象領域のキーワードであるのかというキーワード情報をデータベース108に保持しておく。つまり、例えば、図、表の領域番号403と、キーワード領域の領域番号404とをペアにした情報データベース108に格納される。

【0026】次に、文字認識手段104は、各文字領域から文字行を切り出し、文字行から各文字を切り出し、文字認識処理を行う（ステップ204）。次に、リンク

6

生成手段105は、ステップ203で求めたキーワードと、キーワード以外の文字領域（以下、本文）の関連箇所の間にリンクを生成する（ステップ205）。まず、キーワード情報を基にキーワードを求め、該キーワードに含まれる図などの番号を、文字認識結果から求める。例えば、「図1」、「第1図」などの図番号である。ないで、本文中の同一表現箇所を文字認識結果から検出する。該当する箇所が検出されたら、当該箇所とキーワードとの間にリンクを生成する。

【0027】続いて、リンク生成手段105は、図、表などと本文の間の直接的なリンクを生成する（ステップ206）。ところで、図、表などが上記したように必ずキーワードを持つとは限らない。そこで、本文中の「上図」、「左表」などのように、図、表を表現し、なおかつレイアウト的な情報（上、左などの位置的表現）を含む表現を検出する。本文中の上記した表現の位置と、近傍の図表などの位置は、ステップ202およびステップ204で得た領域の位置情報や文字切り出し、認識結果の位置情報として得られる。したがって、本文中の上記した表現位置から見て、「上」、「下」、あるいは「左」などで、最も近傍にある図や表などの該当する領域を検出し、これら図、表などと本文の該当箇所との間にリンクを生成する（ステップ206）。最後に、構造表現手段106は上記処理の結果を構造化表現する（ステップ207）。リンク関係を表現できる形式としては、例えばHTMLがある。これらリンク情報やキーワード情報など抽出した情報を上記した形式に変換する。

【0028】なお、本発明は上記したものに限定されず、ソフトウェアによっても実現することができ、本発明をソフトウェアによって実現する場合では、図7に示すように、CPU、ROM、RAM、表示装置、ハードディスク、キーボード、CD-ROMドライブ、モデムなどからなる汎用の処理装置を用意し、CD-ROMなどのコンピュータ記憶媒体には、本発明の文書画像の処理機能を実現するプログラムが記録されている。

【0029】ハードディスクなどには、本発明の文書画像処理機能によって、入力文書画像がハイパーテキスト化されて格納されているので、その文書画像を読み出した「図3」をアクセスでリンクさせることにより、「図3」にリンクされた図が表示される。また、同様に、例えば「上図」をクリックすると、「上図」にリンクされた図が表示される。

【0030】【発明の効果】以上、説明したように、本発明によれば、入力文書画像をハイパーテキストで表現して利用するシステムにおいて、文書画像に含まれる図、表に関連するキーワードがある場合には、キーワード内の文字列と本文中の関連箇所との間に自動的にリンクを生成

50

(5)

特開平10-228473

(6)

特開平10-228473

することができ、また、キャラクターがない場合には、図、表とその図、表に関連する本文中の箇所との間に自動的にリンクを生成することができる。

【図面の簡単な説明】

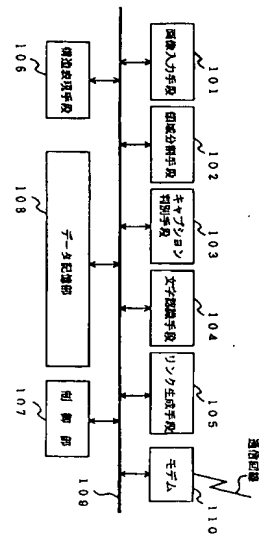
- 【図1】 本発明の実施例の構成を示す。
- 【図2】 本発明の処理フローチャートを示す。
- 【図3】 ステップ203のキャラクター判別処理の詳細なフローチャートを示す。
- 【図4】 文書画像が領域分割された例を示す。
- 【図5】 ステップ303の詳細なフローチャートを示す。
- 【図6】 ステップ304の詳細なフローチャートを示す。

【図7】 本発明をソフトウェアによって実現する場合の構成例を示す。

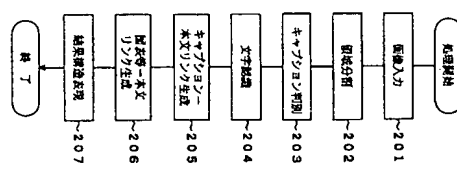
【符号の説明】

- 101 画像入力手段
- 102 領域分割手段
- 103 キャラクション判別手段
- 104 文字認識手段
- 105 リンク生成手段
- 106 構造表現手段
- 107 制御部
- 108 データ記憶部
- 109 データ通信路
- 110 モデム

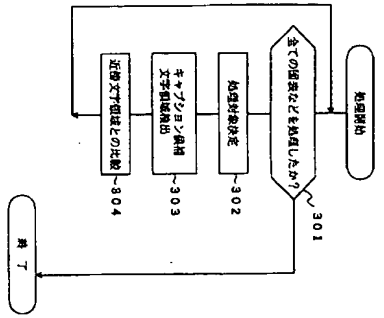
【図1】



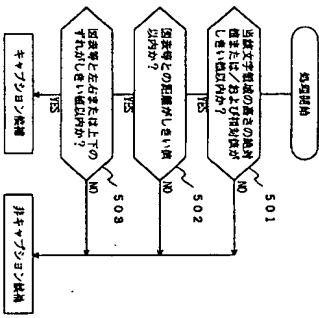
【図2】



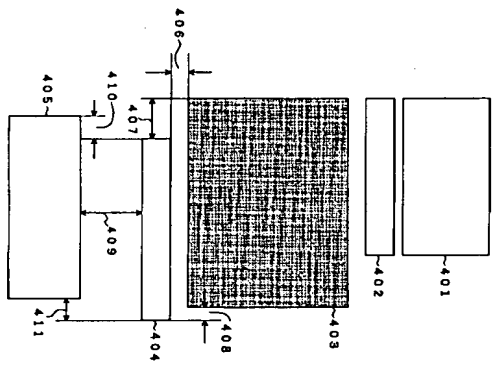
【図3】



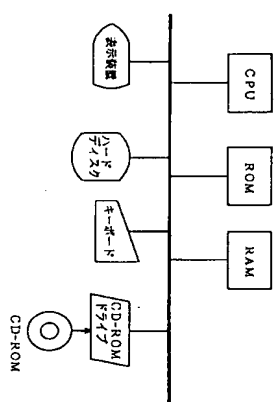
【図5】



【図4】



【図7】



【図6】

